CSCI 4961/6961: Homework 3

Assigned Monday September 28 2020. Due by 11:59pm Monday October 5 2020.

1. (Exercise in proving concavity) Establish the concavity of the geometric mean $f(\mathbf{x}) = \left(\prod_{i=1}^d x_i\right)^{1/d}$ on $\mathbb{R}^d_{++}$ (the set of d-dimensional vectors whose entries are all strictly greater than zero). One potential argument is very similar to the one we gave in class for the convexity of logsumexp (there are other valid arguments). Plot the geometric mean on $[0.1, 3] \times [0.1, 3]$.

2. (Example application of Jensen's inequality) Let $\mathbf{g}$ be a $\mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_d)$ Gaussian random vector. Directly computing the expected length of $\mathbf{g}$ is difficult.

   (a) Prove that the function $\sqrt{x}$ is concave on the interval $(0, \infty)$.

   (b) Compute $\mathbb{E}\left[\|\mathbf{g}\|_2^2\right]$ directly.

   (c) Explain how to use Jensen's inequality to obtain the following upper bound on the expected length of a d-dimensional Gaussian vector:

   $$\mathbb{E}[\|\mathbf{g}\|_2] \leq \sigma\sqrt{d}.$$

   (d) Empirically evaluate the tightness of this bound by using the law of large numbers. Let $\sigma = 1$ and for each integer d in $[50]$, average over 5000 draws to estimate $\mu_d = \mathbb{E}[\|\mathbf{g}\|_2]$. Plot $|\mu_d - \sqrt{d}|/\mu_d$ for these values of d. What do you notice?

3. (Practice with optimality conditions, and projections onto convex sets) Let $\mathcal{B}_2(r)$ denote the Euclidean ball of radius r, centered at $\mathbf{0}$. Show that for any point $\mathbf{x} \notin \mathcal{B}_2(r)$, the projection onto the ball is given by $\mathbf{P}_{\mathcal{B}_2(r)}(\mathbf{x}) = \frac{r}{\|\mathbf{x}\|_2}\mathbf{x}$. That is, show that

   $$\frac{r}{\|\mathbf{x}\|_2}\mathbf{x} = \mathbf{P}_{\mathcal{B}_2(r)}(\mathbf{x}) := \text{argmin}_{\|\mathbf{z}\|_2 \leq r} \|\mathbf{x} - \mathbf{z}\|_2^2.$$

4. (CSCI 6961 students) Consider the ℓ_1 penalized least squares problem

   $$\min_{\mathbf{x} \in \mathbb{R}^d} \frac{1}{2}\|\mathbf{x} - \mathbf{b}\|_2^2 + \lambda\|\mathbf{x}\|_1.$$

   This is the Lasso problem with the identity feature matrix. Argue that the coordinates of the minimizer $\mathbf{x}_\star$ can be solved for independently, then use the optimality conditions to show that the solution of this problem is $\mathbf{x}_\star = S_\lambda(\mathbf{b})$, where S_λ is the *soft-thresholding operator*, defined as

   $$[S_\lambda(\mathbf{b})]_i = \begin{cases} b_i - \lambda & \text{if } b_i \in [\lambda, \infty) \\ 0 & \text{if } b_i \in [-\lambda, \lambda] \\ y_i + \lambda & \text{if } b_i \in (-\infty, -\lambda] \end{cases}$$